Original

Evaluation of the Reliability Coefficient for the Achievement Test of Selected Defensive and Offensive Basketball Skills Using the Classical Theory and Rasch Model

(1) Prof. Dr. Dhurgham Jasim Mohammed Al-Nuaimi

- Mosul University, College of Physical Education and Sports Sciences

Dhurghamjasim@uomosul.edu.iq

(2) Ali Mohammed Ahmed Mulla Al-Zakholi

- Zakho University, College of Education, Department of Physical Education and Sports Sciences

.**Email:** Ali.al-zakholi@uoz.edu.krd

- Mosul University, College of Physical Education and Sports Sciences

ali.21ssp11@student.uomosul.edu.iq

Abstract

The aim of this research is to evaluate the reliability coefficient of the achievement test for selected defensive and offensive basketball skills using both the Classical Test Theory (CTT) and the Rasch Model (Item Response Theory - IRT) as part of modern measurement theory. This research gains its significance by highlighting the effectiveness and accuracy of tools and methods used in measuring athletic performance, contributing to the improvement and development of achievement tests.

The research population was determined to include students from the College of Physical Education and Sport Sciences at the Universities of Duhok and Zakho, with a total study sample of 213 students. Data were collected using scientifically designed measurement tools, and reliability was analyzed using two methods: the classical theory and the Rasch Model.

The results showed that the reliability coefficient based on the classical theory reached 0.85 using the Spearman-Brown method, reflecting a high level of consistency. On the other hand, the Rasch Model's Person-Item Reliability coefficient was 0.712, with fit indices (Infit and Outfit) ranging between 0.71 and 0.87, indicating good item fit to the model.

The comparison between the two methods revealed that while the classical theory provides high overall reliability values, it lacks detailed analysis of item performance. Conversely, the Rasch Model offers a more comprehensive and precise evaluation of item quality and the overall test.

The study recommends revising items with lower performance for improvement and combining both methods for a holistic assessment. It also calls for expanding the sample size and applying the research to other sports fields to enhance the generalizability of the results and the reliability of achievement tests.

Keywords: Reliability, Classical Test Theory, Rasch Model, Defensive and Offensive Skills, Basketball, Achievement Tests.

1- Introduction to the Research:

1.1 Introduction and Significance of the Research:

Basketball is one of the most prominent team sports that combines speed and precision, requiring players to achieve a high level of skillful and tactical performance for success. Athletic performance relies not only on physical abilities but also on fundamental skills, such as defensive and offensive skills, making the accurate measurement of these skills an essential necessity to ensure the improvement of both individual and team performance.

Achievement tests are an effective means of evaluating these skills, as they provide objective data that contribute to assessing players' progress. However, the results of these tests can only be reliable if they demonstrate a high level of stability (reliability), which is considered one of the key indicators of the quality of any test. Therefore, it has become imperative to apply advanced analytical methods to ensure the accuracy and reliability of these tests.

The significance of this research lies in analyzing the reliability coefficient of achievement tests for basketball's defensive and offensive skills by comparing them using the Classical Test Theory (CTT) and the Rasch Model, which is one of the most notable models in the framework of modern measurement theory. This comparison aims to provide a comprehensive scientific perspective on the best methods for analyzing and improving sports tests to serve training and evaluation objectives.

This research holds great importance in the sports field due to the pivotal role that achievement tests play in assessing and developing performance. It seeks to improve the process of measuring fundamental basketball skills by offering a detailed analysis of the reliability coefficient using various methods that combine traditional and modern approaches. This not only enhances the accuracy of the tests but also contributes to improving training strategies, positively reflecting on players' performance in competitions.

Additionally, this research represents a qualitative addition to the academic field by shedding light on the Rasch Model as a modern tool for data analysis. This enhances the reliability of results and paves the way for future studies that deepen the understanding of the relationship between classical and modern theories in sports performance analysis. Thus, this research aims to support efforts to develop sports tests and ensure their use as an effective scientific tool in the sports field.

1.2 Research Problem:

Achievement tests are a fundamental tool for evaluating athletic skills, as they provide quantitative data that contribute to measuring and analyzing performance levels. However, ensuring the accuracy and reliability of these tests remains a persistent challenge for researchers and coaches alike, especially since many classical tests rely on measurement methods that may lack precision in certain aspects.

In basketball, defensive and offensive skills demand a high level of performance that must be measured accurately to ensure the development of effective training strategies. However, relying solely on the Classical Test Theory (CTT) in test analysis may not fully reflect the nature of players' performance and its variability, as it primarily focuses on overall traits without considering individual factors. Conversely, the Rasch Model within modern measurement theory offers a precise analytical tool that takes into account both individual characteristics and the test itself, enhancing the reliability and accuracy of the results.

The research problem lies in the need to evaluate and analyze the reliability coefficient of achievement tests for selected defensive and offensive basketball skills and to compare the effectiveness of the Classical Test Theory and the Rasch Model in providing more accurate and objective results. This study aims to address the following key question:

- How effective are the Classical Test Theory and the Rasch Model in analyzing the reliability coefficient of achievement tests for defensive and offensive basketball skills?

1.3 Research Domains:

- **Temporal Domain:** From October 1, 2021, to November 25, 2024.
- **Human Domain:** Students of the College of Physical Education and Sport Sciences at the University of Duhok and students of the Department of Physical Education and Sport Sciences at the College of Education, University of Zakho.
- **Spatial Domain:** The College of Physical Education and Sport Sciences at the University of Duhok and the Department of Physical Education and Sport Sciences at the College of Education, University of Zakho.

1.5 Definition of Terms

- **Achievement Test:** The achievement test is defined as an organized tool aimed at measuring the level of achievement of a sample from a specific population in terms of knowledge and skills related to a particular academic subject or sport (game) they have previously learned formally. This is achieved through their responses to a set of selected questions (items) representing the content of that theoretical-academic subject. (Bharadway, 2024, 135)
- **Defensive Skills in Basketball:** Defensive skills in basketball are a set of abilities possessed by a basketball player and performed during a game while in a defensive position, within the legal framework of the game. These skills can be executed individually, collectively, or as a team, aiming to prevent the opposing team from scoring a successful basket (point or multiple points) against their team. Key defensive skills in basketball include readiness stance, foot movement,

interception of passes, defending against the shooter, defending against the dribbler, guarding the pivot player, and collecting rebound balls. (Crudeli, 2023, 26)

- **Offensive Skills in Basketball:** Offensive skills in basketball are a set of abilities possessed by a basketball player and performed during a game while in an offensive position, meaning when their team is in possession of the ball. These skills aim to score as many points as possible in the opposing team's basket to ensure their team's victory in the game. Offensive skills include shooting, passing, and dribbling. (Oliver, 2024, 53)
- **Classical Test Theory (CTT):** This theory assumes in the field of measurement that the variance of measurement errors is equal for all individuals taking the test, regardless of their varying abilities or inconsistencies in their performance. (Clauser and Bunch, 2021, 113)

2. Research Procedures

2.1 Research Methodology:

The researchers adopted the descriptive method with a comparative approach due to its suitability for the nature of the current study.

2.2 Research Population and Sample:

The research population consists of students from the College of Physical Education and Sport Sciences at the University of Duhok and the Department of Physical Education and Sport Sciences at the College of Education, University of Zakho, during the academic year (2023–2024). The total number of students is 213.

2.3 Data Collection Methods and Tools:

1. **Content Analysis:** To obtain accurate and comprehensive data about the research variables, a systematic analysis of content related to scientific sources and studies was conducted. These sources were carefully selected to ensure coverage of both theoretical and practical aspects of the topic, enhancing the accuracy and credibility of the findings.
2. **Personal Interviews:** Personal interviews, some conducted via phone, were held with a group of experts and specialists in the fields of measurement and evaluation, teaching methods, and basketball. These interviews included research questions related to the study methodology and the application of the tests based on the Classical Test Theory (CTT) and the Rasch Model.
3. **Questionnaire:**
   The researchers prepared a scale comprising 43 items, which was then reviewed by a group of experts and specialists to ensure its suitability. After revisions, the questionnaire was administered to a construction sample of 106 students.
   - In the Classical Test Theory (CTT) framework, statistical analyses showed that 30 items were valid.

  - In the Rasch Model framework, statistical analyses indicated that 27 items were valid.

The steps for constructing the achievement test were implemented based on the Classical Test Theory and the Rasch Model within the framework of modern measurement theory.

2. Research Procedures

2.1 Research Methodology: The researchers adopted the descriptive method with a comparative approach due to its suitability for the nature of the current study.

2.2 Research Population and Sample: The research population consists of students from the College of Physical Education and Sport Sciences at the University of Duhok and the Department of Physical Education and Sport Sciences at the College of Education, University of Zakho, during the academic year (2023–2024). The total number of students is 213.

2.3 Data Collection Methods and Tools:

1. **Content Analysis:** To obtain accurate and comprehensive data about the research variables, a systematic analysis of content related to scientific sources and studies was conducted. These sources were carefully selected to ensure coverage of both theoretical and practical aspects of the topic, enhancing the accuracy and credibility of the findings.
2. **Personal Interviews:** Personal interviews, some conducted via phone, were held with a group of experts and specialists in the fields of measurement and evaluation, teaching methods, and basketball. These interviews included research questions related to the study methodology and the application of the tests based on the Classical Test Theory (CTT) and the Rasch Model.
3. **Questionnaire:** The researchers prepared a scale comprising 43 items, which was then reviewed by a group of experts and specialists to ensure its suitability. After revisions, the questionnaire was administered to a construction sample of 106 students.
   - In the Classical Test Theory (CTT) framework, statistical analyses showed that 30 items were valid.
   - In the Rasch Model framework, statistical analyses indicated that 27 items were valid.

The steps for constructing the achievement test were implemented based on the Classical Test Theory and the Rasch Model within the framework of modern measurement theory.

2.4.2 Reliability:

Reliability of the Achievement Test for Selected Defensive and Offensive Basketball Skills According to the Classical Test Theory:

1. **Reliability Using the Split-Half Method:**

The reliability of the test was calculated using the split-half method, which involves dividing the test items into two equal or equivalent parts and then measuring the correlation between the results of these two parts.
The split-half method for measuring reliability entails dividing the test items into two groups: one comprising the odd-numbered items and the other the even-numbered items. After dividing the items, the correlation coefficient between the results of the two groups is calculated. The value obtained from this analysis serves as an indicator of the test's reliability. However, this value requires adjustment using the Spearman-Brown formula to achieve a more accurate estimation of the overall reliability of the test. (Parihar and Kmar, 2023, p. 69)
The researchers calculated the reliability of the test using the split-half method. The test was administered on March 27, 2024, to a sample of 30 students. The data were then divided into odd-numbered and even-numbered items, and the correlation coefficient between the two parts was determined, as shown in Table (1).

Table (1): Split-Half Method Reliability Coefficient

Reliability Method	Reliability Value (Pearson)	Reliability Value (Spearman-Brown)
Reliability Using the Split-Half Method	0.74	0.85

From Table (1), it is evident that the reliability coefficients for the scale are high and significant. The correlation values used to calculate the reliability coefficients are considered strong when they fall within the range of 0.70 to 0.95 (Swales and others, 2023, p. 36).

In this study, the Spearman-Brown adjusted reliability coefficient is 0.85, which lies within the upper range of acceptable reliability, indicating a high level of consistency for the test as a whole. Similarly, the Pearson correlation value of 0.74 also falls within the acceptable range, further supporting the robustness and reliability of the scale used in this research.

These results confirm the test's ability to consistently measure the targeted basketball skills, making it a reliable tool for evaluating defensive and offensive performance.

- **Reliability of the Achievement Test for Selected Defensive and Offensive Basketball Skills According to Modern Measurement Theory (Rasch Model):**

1. **Reliability Coefficient Using the Person-Item Reliability Index:**

The Person-Item Reliability Index was calculated, which represents the ratio of variance attributed to individuals to the variance attributed to error. A high value for this index indicates that the model provides an accurate estimation of abilities.

In this study, the reliability coefficient using the Person-Item Reliability method was approximately 0.71. This value suggests that the Rasch Model estimates abilities with reasonable accuracy, reflecting good differentiation between individuals and error variance.

These findings confirm the effectiveness of the Rasch Model in providing precise measurements of the targeted basketball skills. The results are visually represented in Figure (1).

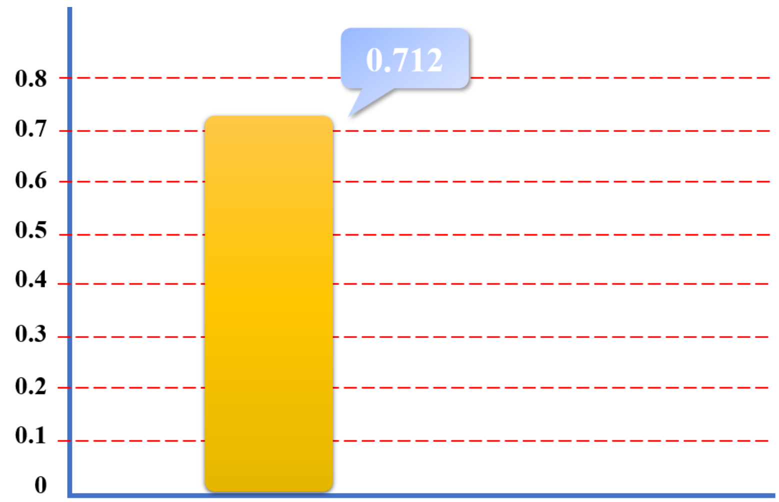

Figure (1) illustrates the reliability coefficient using the Person-Item Reliability index.

The Item Fit Statistics, which include Infit and Outfit indices for different items in the model, were calculated. These indices provide an evaluation of how well each item fits the Rasch Model:

- Outfit Statistics values range approximately between 0.71 and 0.84.
- Infit Statistics values range approximately between 0.72 and 0.87.

These values indicate that the majority of the items demonstrate good fit within the model, ensuring accurate measurement of the intended skills. The findings are presented in Table (2) and visually represented in Figure (2).

Table (2) shows the reliability coefficient using Item Fit Statistics.

S	Infit Statistics	Outfit Statistics
1	0.86	0.72
2	0.80	0.72
3	0.78	0.74
4	0.86	0.72
5	0.83	0.83
6	0.83	0.74
7	0.80	0.84

8	0.76	0.80
9	0.77	0.82
10	0.82	0.82
11	0.73	0.77
12	0.83	0.77
13	0.84	0.77
14	0.77	0.78
15	0.85	0.82
16	0.81	0.78
17	0.80	0.78
18	0.85	0.75
19	0.85	0.82
20	0.83	0.73
21	0.87	0.80
22	0.84	0.80
23	0.86	0.71
24	0.84	0.81
25	0.86	0.78
26	0.77	0.73
27	0.81	0.80

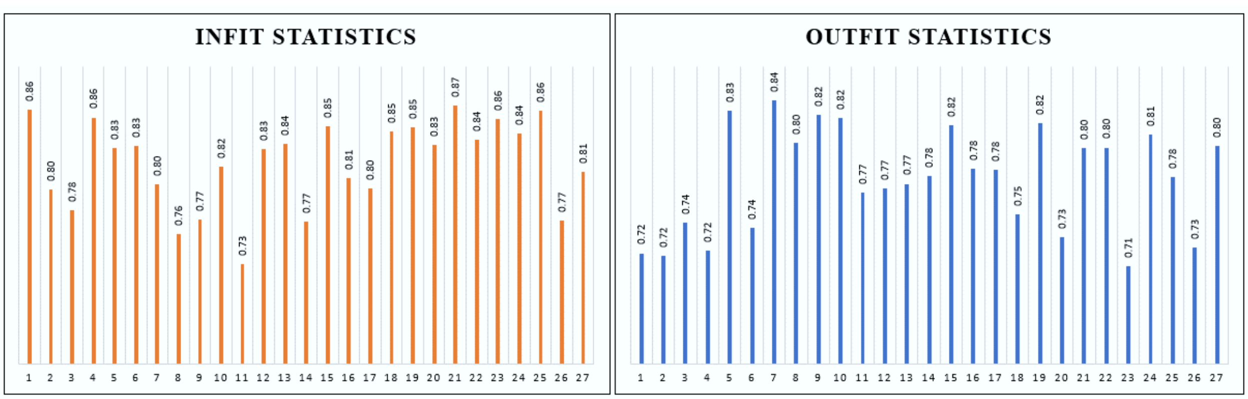

Figure (2) illustrates the reliability coefficient using Item Fit Statistics.

The left side of the figure represents the Infit Statistics, which indicate the alignment of items with the Rasch Model, particularly for items of medium difficulty. The ideal values for Infit are close to 1. The values shown in the graph range between 0.71 and 0.87, suggesting good item fit with some variability.

The right side of the figure represents the Random Adjusted Outfit Statistics, which reflect the alignment of items with the model, focusing on items with very high or very low difficulty levels. The ideal values for Outfit are also close to 1. The displayed values range between 0.71 and 0.87, indicating good item fit with some variability.

The values shown in the graph provide a positive indication of item alignment with the model, as most values are close to the ideal range. This reinforces confidence in the reliability of the data and the accuracy of performance interpretation.

2.4.3 Objectivity:

The current research relies on data collection methods that ensure objectivity and eliminate personal bias from the researchers. Instead of depending on the researchers' subjective evaluation, the data is recorded through the participants' direct responses to the questionnaire items.

This approach enhances objectivity by relying entirely on the straightforward and honest feedback of the individuals in the study sample. It ensures that the extracted data accurately reflects their opinions and actual experiences. By adopting this systematic and independent method of data collection, the study achieves a high level of objectivity, free from the researchers' personal influence.

This emphasis on objectivity strengthens the reliability and scientific validity of the results, ensuring they are credible and unbiased.

2.5 Research Experiments

1. Pilot Experiment:

The pilot experiment is one of the most critical procedures necessary to ensure the accuracy of scientific work. This experiment aims to overcome difficulties and save both material and physical effort by testing the scale before finalizing it. After completing all scientific and administrative preparations, the researchers conducted a pilot experiment on March 25, 2024, with a sample of 10 students. The objectives of this experiment were as follows:

1. To assess the sample's ability to comprehend the purpose of the scale.
2. To evaluate the clarity of the items.
3. To ensure the clarity of the test instructions.
4. To determine the time required to complete the scale.
5. To identify challenges or issues that might face the researchers.
6. To address any questions or concerns raised by the participants.
7. To evaluate the sample's responsiveness to the researchers.
8. To observe the sample's performance in the tests.
9. To organize the supporting research team (Appendix 8).

The results of the pilot experiment demonstrated that the scale was ready for implementation in the main experiment, with no significant issues or obstacles identified. This allowed the researchers to proceed with the next steps in developing the scale.

3.7 Statistical Methods Used

The following statistical methods and tools were employed to analyze the data in this research:

- Percentage.
- Ease Index.
- Difficulty Index.
- Item Discrimination Index.
- Internal Consistency.
- Factor Analysis.
- Independence of Estimating Individuals' Abilities:
    - Internal Fit Statistics (INFIT) using Mean Squares (MNSQ).
    - Internal Fit Statistics (INFIT) using Z-Statistic (ZSTD).
    - External Fit Statistics (OUTFIT) using Mean Squares (MNSQ).
    - External Fit Statistics (OUTFIT) using Z-Statistic (ZSTD).
- Independence of Estimating Item Difficulty:
    - Internal Fit Statistics (INFIT) using Mean Squares (MNSQ).
    - Internal Fit Statistics (INFIT) using Z-Statistic (ZSTD).
    - External Fit Statistics (OUTFIT) using Mean Squares (MNSQ).
    - External Fit Statistics (OUTFIT) using Z-Statistic (ZSTD).
- Parallelism of Item Characteristic Curves (Item-Characteristic):
    - Item Parameter (b).
    - Log-Likelihood Probability.
- Correlation (Reliability Coefficient).
- Spearman-Brown Formula.
- Reliability Coefficient Using Person-Item Reliability.
- Reliability Coefficient Using Item-Fit Reliability.

Statistical Software:

- SPSS (Statistical Package for the Social Sciences).
- Microsoft Excel.
- WINSTEPS® Rasch Analysis and Rasch Measurement Software for processing data related to the Rasch Model (dichotomous, binary, and multiple-choice items).
- 3BILOG-MG software for processing data under modern measurement theory (latent trait theory using the Rasch Model).

3.1 Discussion of Results: Reliability According to Classical Test Theory

The results obtained from the reliability analysis using Classical Test Theory (CTT) indicate that the scale used possesses an acceptable level of reliability based on two primary methods: split-half reliability and test-retest reliability.

- **Split-Half Reliability:**

Through the analysis of split-half reliability results using two different approaches—Pearson correlation coefficient and Spearman-Brown coefficient—the extracted values fall within the scientifically acceptable range for reliability.

1. The reliability coefficient using the Spearman-Brown method recorded a value of 0.85, a high value that indicates a strong correlation between the two equal halves of the test.
2. The Pearson correlation coefficient recorded a value of 0.74, which is relatively lower but still falls within the acceptable range (0.70–0.95). This supports the reliability of the test when employed in practical applications.

- **Interpretation of the Results:**

These results demonstrate that the scale achieves a high level of reliability. The high values suggest strong internal consistency among the items composing the scale, which is a positive indicator of the quality of the test design.

However, it is noteworthy that there is a slight variation between the two coefficients (Pearson and Spearman-Brown), reflecting the influence of the method used to calculate reliability on the final results. This variation highlights the importance of using multiple approaches to assess reliability to obtain a comprehensive understanding of the scale's consistency.

Overall, the findings support the effectiveness of the scale in consistently measuring the targeted basketball skills, providing a strong foundation for its use in similar applications.

3.2.1 Discussion of Results: Reliability Using Person-Item Reliability in the Rasch Model:

The results obtained through the Person-Item Reliability Index indicate that the reliability coefficient was 0.712, which is considered acceptable according to the theoretical framework of reliability indicators in the Rasch Model.

This value represents the ratio of true variance to total variance in the data, demonstrating that the Rasch Model has a good capability to accurately estimate individual abilities (persons) and item difficulty.

The reliability value of 0.712 indicates a relatively low error variance, reflecting the test's reliability in achieving stable and consistent measurements. This confirms the scale's effectiveness in providing precise assessments of the targeted basketball skills while maintaining alignment with individual differences and item difficulty levels.

The findings underscore the utility of the Rasch Model in producing reliable results, ensuring that the measurement process is both objective and consistent across various skill levels.

3.2.2 Discussion of Results: Reliability Using Item Fit Statistics

The analysis of **Item Fit Statistics**, both **Infit** (internal fit) and **Outfit** (external fit), reveals values ranging between **0.71 and 0.87**, indicating an overall good fit. These values reflect the degree to which the data aligns with the requirements of the Rasch Model, demonstrating that most items fit the model well.

- **Infit Statistics:** These values assess the alignment of items with the model, particularly focusing on items with medium difficulty. Ideal values are close to **1**. The measured values, ranging between **0.72 and 0.87**, demonstrate a good level of fit, suggesting that the items are well-designed to accommodate various skill levels and align with the majority of participants' abilities.
- **Outfit Statistics:** These indicators evaluate the external fit of items, with a focus on items with very high or very low difficulty levels. The recorded values, ranging between **0.71 and 0.84**, also reflect a good level of fit, further enhancing the confidence in the reliability of the test.

- **Comparison of Reliability Results Between Classical Test Theory and the Rasch Model**

1. Level of Reliability in Both Models:

- According to Classical Test Theory (CTT), the results indicated a high level of reliability, with a test-retest and Spearman-Brown coefficient of 0.85. This value falls within the scientifically acceptable high range (0.70–0.95), reflecting a strong level of consistency when the test is applied under similar conditions.
- In contrast, the reliability value for the Rasch Model (Item Response Theory - IRT) was 0.712, based on the Person-Item Reliability index. While this value is considered scientifically acceptable, it is lower than the reliability obtained using CTT.

2. Statistical Interpretation:

- CTT relies on the analysis of total scores without accounting for variations between items or individual performances. This approach may inflate reliability values in cases where there is general consistency in performance.
- The Rasch Model, on the other hand, provides a more detailed analysis, evaluating each item's fit and its ability to differentiate between individuals. Consequently, reliability values may be slightly lower but reflect a more accurate measurement of the targeted traits.

3. Item Fit:

- Results from CTT show that the scale demonstrates a general level of consistency but do not provide insights into the performance of individual items. This limitation may lead to the inclusion of poorly performing items without understanding their impact.
- The Rasch Model, through Infit and Outfit Statistics, showed that most items fall within the ideal range of 0.71–0.87. These statistics offer a detailed analysis of item quality, enabling improvements by revising or removing poorly fitting items.

4. Purpose of Reliability in Each Model:

- CTT aims to achieve high overall reliability, reflecting the stability of total scores when the test is repeated. This makes it suitable for general purposes that do not require detailed analysis of individual performance.
- The Rasch Model focuses on providing precise and objective measurement of the targeted traits, aiming to reduce the influence of external factors (e.g., item difficulty or bias), making it more appropriate for studies seeking to enhance the quality of measurement tools.

5. Generalizability:

- Reliability results from CTT are often tied to the specific sample on which the test is administered, meaning that the results may not be generalizable to other samples.
- In contrast, the Rasch Model is known for its ability to produce results that are more generalizable, as its analysis is independent of sample characteristics. This enhances the reliability of the test when applied in different settings or with new samples.

6. Individual Differences:

- In CTT, high reliability values are used as a general indicator of test quality, making it easier to apply without frequent modifications.
- Conversely, the Rasch Model emphasizes the need for detailed evaluation and analysis of each item individually, making it a more efficient tool for improving tests and ensuring greater fairness and objectivity.

Summary of the Comparison: The comparison reveals that Classical Test Theory provides high reliability values that are easy to interpret but lacks the detailed analysis of item quality and individual performance. On the other hand, the Rasch Model, while yielding slightly lower reliability values, offers a deeper and more precise analysis, making it a more advanced tool for developing and improving achievement tests.

Conclusion: Combining the two approaches can provide a comprehensive evaluation. CTT can be used to assess overall reliability, while the Rasch Model can analyze items in detail and improve the quality of tests continuously. This integration ensures both general consistency and precise, objective measurement.

4. Conclusions and Recommendations

4.1 Conclusions:

4 **Reliability of the Test According to Classical Test Theory:** The results of the analysis using Classical Test Theory (CTT) revealed high reliability values, with the reliability coefficient calculated using the Spearman-Brown method reaching **0.85**. This indicates a high level of consistency and reliability when the test is applied in similar settings. However, CTT demonstrates limitations in providing detailed insights into the performance of individual items.

5 **Reliability of the Test According to the Rasch Model:** The reliability results according to the Rasch Model showed a Person-Item Reliability value of 0.712, which is considered acceptable and reflects the model's ability to accurately measure the targeted traits. Additionally, the item fit statistics (Infit and Outfit) were within the ideal range of 0.71–0.87, indicating the quality of item design and their alignment with the model.

6 **Comparison Between the Two Theories:** Although Classical Test Theory demonstrated higher overall reliability values, the Rasch Model provided a more detailed and accurate analysis, highlighting the strengths and weaknesses of the test. The Rasch Model proved to be a more flexible and objective tool for data analysis.

7 **Importance of Combining the Two Approaches:** The findings emphasize the value of using both approaches to achieve a comprehensive evaluation. While CTT can be utilized to assess overall reliability, the Rasch Model can contribute to improving test quality through detailed analysis of item performance.

4-2 Recommendations

- **Improving Item Quality:** Based on the results of the Rasch Model, it is recommended to review items with lower fit values (Infit and Outfit) and work on modifying or replacing them to ensure improved test accuracy and reliability.
- **Integrating Both Theories in Data Analysis:** It is advised to use Classical Test Theory as a first step to evaluate overall reliability and then employ the Rasch Model to analyze detailed item performance. This approach contributes to creating more accurate and equitable tests.
- **Expanding the Sample:** Applying the test to larger and more diverse samples is recommended to ensure the generalizability of results and enhance reliability.
- **Developing Achievement Tests:** The findings of this study should be utilized in the development of new tests that incorporate the principles of the Rasch Model, especially in fields requiring high levels of accuracy and objectivity.

- **Training on the Rasch Model:** Workshops and training programs should be provided for researchers and trainers on using the Rasch Model in test analysis to improve their capabilities in refining measurement tools.
- **Conducting Future Research:** Future studies are recommended to apply the Rasch Model to achievement tests in other sports and educational fields, contributing to the overall improvement of measurement quality and evaluation practices.

References

1. Bharadway, Anand (2024): **Checkpoints VCE General Achievement Test (GAT)**, Cambridge University Press. UK.
2. Crudeli, Christian (2023): **My basketball DEVELOPMENT FUNDAMENTALS**, Grupo Editorial circulo Rojo SL, Spain.
3. Oliver, Dean (2024**): Basketball beyond Paper: Insights into the Game's Analytics Revolution,** University of Nebraska Press, USA.
4. Clauser, Brian E. and Bunch, Michael B. (2021): **The History of Educational Measurement**, published by Routledge. UK.
5. David, Andrich and Marais, Ida (2019): **A Course in Rasch Measurement Theory Measuring in the Educational, Social and Health Sciences,** SpringerLink, UK.
6. Parihar, Amarjeet and Kumar, Sushi (2023): **MEASUREMENT AND EVALUATION IN EDUCATION**, RLAL Education Punlishers, India.
7. Swales, Bridgitte and others (2023): **The reliability and suitability of strength assessments in frail and pre-frail older adults: recommendations for strength testing in older populations**, BMG Geriatrics, United Kingdom.